Stochastic optimization algorithms
Lecture 3, 20180907

Classical optimization methods (ii)

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Today’s learning goals

- After this lecture you should be able to
  - Describe and use Newton-Raphson’s method
  - Define convex optimization problems
  - Describe and use the method of Lagrange multipliers
  - Describe and use an analytical method for constrained optimization
  - Describe and use the penalty method
  - List and discuss limitations of classical optimization

From last time...

- Iterative algorithms

$$\boldsymbol{x}_{j+1} = \boldsymbol{x}_j + \eta_j \boldsymbol{d}_j$$

pp. 12-14

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Newton-Raphson's method

- Consider the second-order expansion of $f(x)$:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)\ (x - x_0)^2 \equiv f_{[2]}(x)$$

- Minima occur at stationary points, where $f'_{[2]}(x) = 0$:

$$f'_{[2]}(x) = f'(x_0) + \ f''(x_0)(x - x_0) = 0 \Rightarrow x^* = \ x_0 - \frac{f'(x_0)}{f''(x_0)}$$

pp. 21-24

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Newton-Raphson's method

- Thus, in general (iterative method!):

$$x_{j+1} = x_j - \frac{f'(x_j)}{f''(x_j)}$$

- Can be generalized to n dimensions (Newton's method), see p. 22.
- To do (for you!): Example 2.4.

pp. 21-24

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Today's learning goals

- After this lecture you should be able to
  - **Describe and use Newton-Raphson's method** ✓
  - Define convex optimization problems
  - Describe and use the method of Lagrange multipliers
  - Describe and use an analytical method for constrained optimization
  - Describe and use the penalty method
  - List and discuss limitations of classical optimization

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Convex optimization problems

- A convex optimization problem is a special kind of constrained optimization problem, and it occurs if
  - f and g (the inequality constraints) are convex.
  - h (the equality constraints) are *affine*, i.e.

    $$h_i(\boldsymbol{x}) = A_i{}^T \boldsymbol{x} + b_i$$
- In that case S is convex and any local minimum is a global minimum.

pp. 24-25

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Today’s learning goals

- After this lecture you should be able to
  - Describe and use Newton-Raphson’s method ✓
  - **Define convex optimization problems** ✓
  - Describe and use the method of Lagrange multipliers
  - Describe and use an analytical method for constrained optimization
  - Describe and use the penalty method
  - List and discuss limitations of classical optimization

Optimization under constraints

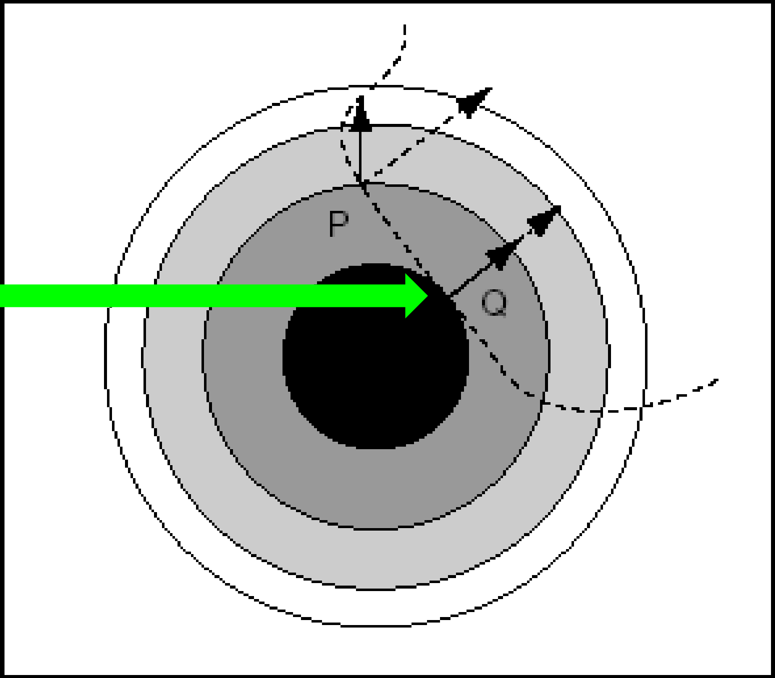
- Specific methods for solving convex optimization problems exist but will not be considered here.
- Instead, we will now consider three general methods for constrained optimization.
- The first two methods enumerate the possible optima, and are applicable when the number of variables is small, whereas the third method (the Penalty method) has more general applicability (and can also be used with stochastic optimization methods).

pp. 25-33

The method of Lagrange multipliers

- Movement along level curves (for a 2D problem – but can be generalized), in case of a single equality constraint:

Local optimum.
Here, $\nabla f = -\lambda \nabla h$



pp. 25-28

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

The method of Lagrange multipliers

- Lagrange multiplier method. Consider the function

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda h(x_1, x_2)$$

- Consider the equation $\nabla L = 0$:

$$\frac{\partial L}{\partial x_1} = \frac{\partial f}{\partial x_1} + \lambda \frac{\partial h}{\partial x_1} = 0$$

$$\frac{\partial L}{\partial x_2} = \frac{\partial f}{\partial x_2} + \lambda \frac{\partial h}{\partial x_2} = 0$$

$$\frac{\partial L}{\partial \lambda} = h = 0$$

pp. 25-28

The method of Lagrange multipliers

- The local optima of f subject to the equality constraint(s) h_i = 0 can thus be found by computing the stationary points of L.
- Note: Finds both minima and maxima!
- To do (for you!): Example 2.6.

- *Note: An exercise class will be held on 20180918. In that lecture, I will demonstrate the solution to several problems involving classical optimization , e.g. problems 2.12 and 2.13, as well as some problems related to EAs.*

pp. 25-28

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Today's learning goals

- After this lecture you should be able to
  - Describe and use Newton-Raphson's method ✓
  - Define convex optimization problems ✓
  - **<u>Describe and use the method of Lagrange multipliers</u>** ✓
  - Describe and use an analytical method for constrained optimization
  - Describe and use the penalty method
  - List and discuss limitations of classical optimization

An analytical method for constr. opt.

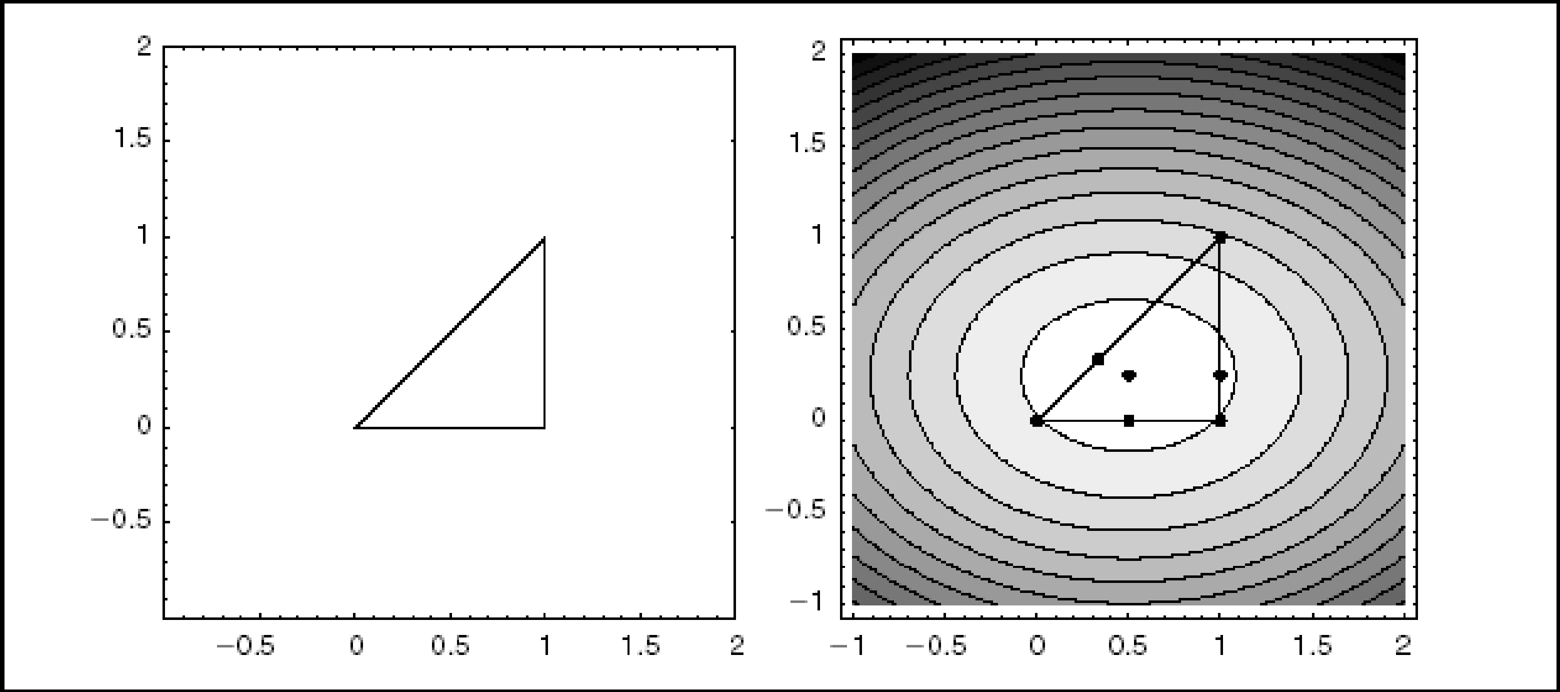
- Method description:
  1. Find the stationary points in the interior of S.
  2. Find the stationary points of the restriction of $f(x)$ to δS.
  3. Investigate the points one by one.
- This method can be used in low-dimensional problems.

pp. 29-30

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

An analytical method for constr. opt.

- Example 2.7:



pp. 29-30

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Today’s learning goals

- After this lecture you should be able to
  - Describe and use Newton-Raphson’s method ✓
  - Define convex optimization problems ✓
  - Describe and use the method of Lagrange multipliers ✓
  - **Describe and use an analytical method for constrained optimization** ✓
  - Describe and use the penalty method
  - List and discuss limitations of classical optimization

The penalty method

- The penalty method transforms a constrained optimization problem to an unconstrained one.
- Consider the *penalty function* p:

$$p(\boldsymbol{x};\mu) = \mu\left(\sum_{i=1}^{m}(\max\{g_i(\boldsymbol{x}),0\})^2 + \sum_{i=1}^{k}(h_i(\boldsymbol{x}))^2\right)$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

The penalty method

- $p \geq 0$, with equality only if all constraints are satisfied.
- Thus, minimizing f subject to g and h is equivalent to minimizing $f_p(\boldsymbol{x};\mu) \equiv f(\boldsymbol{x}) + p(\boldsymbol{x};\mu)$ without constraints *as μ tends to infinity*.
- Numerical approach: Start with a small μ (e.g. 1), find the optima (using e.g. gradient descent or any other method), then increase μ, again find the optima etc. etc.
- As μ gets larger the optima (normally) converge towards those of the constrained problem.
- To do (for you!): Examples 2.8 and 2.9.

Today's learning goals

- After this lecture you should be able to
  - Describe and use Newton-Raphson's method
  - Define convex optimization problems
  - Describe and use the method of Lagrange multipliers
  - Describe and use an analytical method for constrained optimization
  - **Describe and use the penalty method**
  - List and discuss limitations of classical optimization

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

CHALMERS

Limitations of classical optimization

- Classical methods are typically excellent *when they can be applied.*
- However, many kinds of (real-world) problems require other methods. Examples are problems where …
  - … the objective function cannot be specified explicitly as a mathematical function.
  - … the objective function is non-differentiable, and perhaps contains a mixture of (say) Boolean and numerical variables.
  - … the number of *variables* itself varies (for example in optimization of neural network of varying size).

Limitations of classical optimization

- In the rest of the course, we shall consider stochastic optimization methods, which can easily handle problems of the kinds just mentioned.
- **Important!** Make sure that you *attend the next few lectures* (as well as the programming session on Tuesday evening!).
- **Note!** On Tuesday the evening session (18.00) is in MT9, M11-MT13, in the basement of the M building.

Today's learning goals

- After this lecture you should be able to
  - Describe and use Newton-Raphson's method ✓
  - Define convex optimization problems ✓
  - Describe and use the method of Lagrange multipliers ✓
  - Describe and use an analytical method for constrained optimization ✓
  - Describe and use the penalty method ✓
  - **List and discuss limitations of classical optimization** ✓

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Introductory programming problem

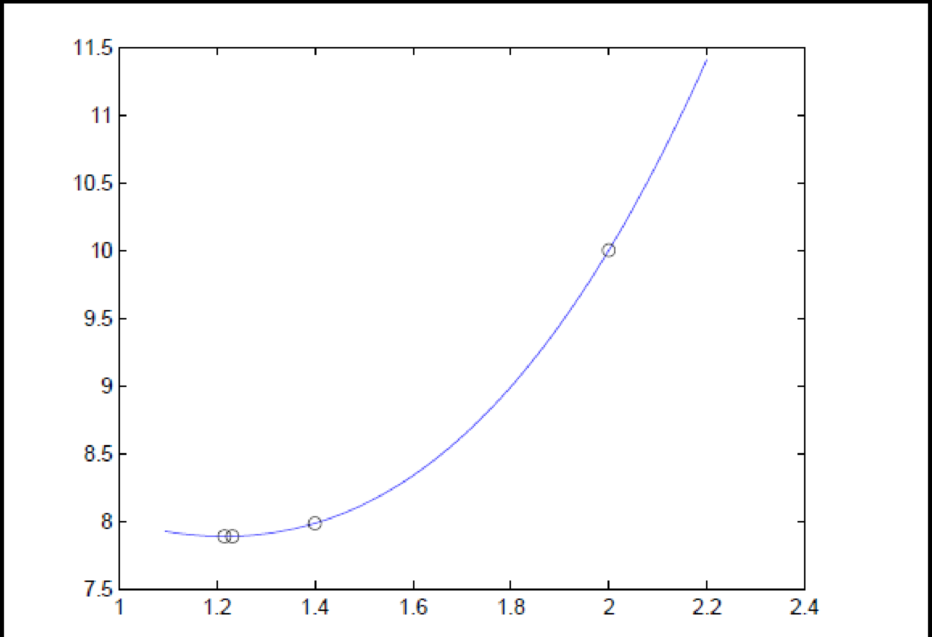
- Mandatory! Must be solved (and handed in, via e-mail) separately by each student.
- Available on the web page.
- Strict deadline (NOTE!) 20180914.
- You will receive feedback (potentially relevant for your work with home problems 1 and 2) as soon as possible.

Introductory programming problem

- Important:
  - The main aim is for you to learn how to write clear, well-structured program code (in Matlab, in this case).
  - Make sure to *follow the coding standard* (available on the course web page!).
  - Before submitting the solution (as an e-mail attachment), check that it can run on the Matlab version available at Chalmers.
  - **NOTE:** Write your full name and civic registration number in the body of the e-mail! (needed for our administation).

Introductory programming problem

- Problem: Implement a Newton-Raphson solver for polynomials, using Matlab.



Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde